2009: A LOOK BACK AT OUTCOME MEASURES

BRUCE G. IWASAKI

Legal Aid Foundation of Los Angeles

submitted for

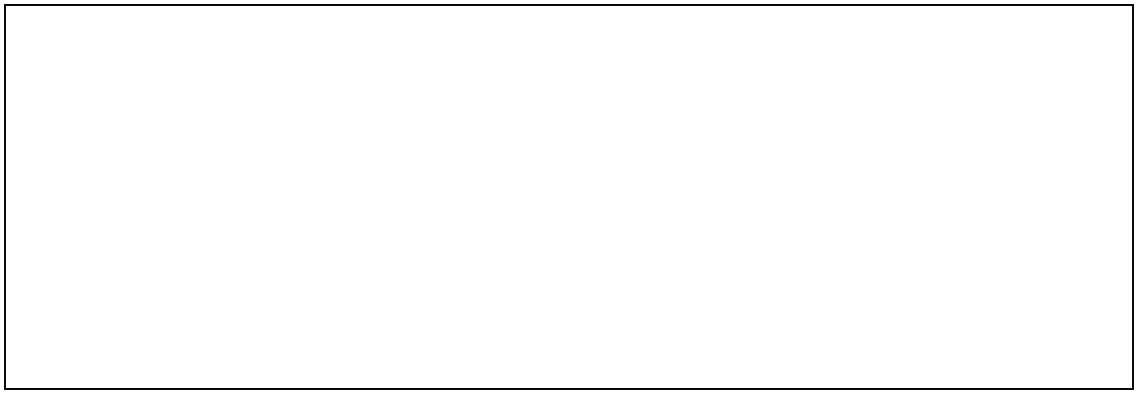
Outcomes Summit II

Cincinnati, Ohio

June 24 - 26, 2004

2009: A Look Back at Outcome Measures

Bruce G. Iwasaki



In her preparation for the 2009 Conference on Outcomes in Legal Services, "Quality and Competition: Should we eat what we kill?", Teresa, the Director of the Legal Aid Foundation of Los Angeles, gathered together both her thoughts and five years of files she had collected on the issue of client outcomes. Even before Teresa became executive director of LAFLA in 2005, the national legal services community had had vigorous discussions on improving the quality of legal services and how quality should be defined, measured and encouraged.

The LAFLA executive director's views were still evolving on the topic. She recognized that quality legal services have objective and measurable aspects – as in a hospital's protocols for practicing quality medical care. But because law is not physiology, and assessing the causes of multivariable results is not physics, Teresa also wondered whether any formula for defining law firm quality was as subjective as determining good art.

Dual role

Beyond issues of definition and application, there were basic issues of purpose. The Legal Services Corporation was like every important political creation: a product of compromise. It thus had to balance competing forces. LSC was responsible to Congress for ensuring its grantees complied with the law. But the legislation also established a system of legally separate, locally governed entities, and jealously insulated the staff lawyer's professional judgment even from those governing boards. During the early discussions on quality, Teresa perceived a struggle between the two contradictory imperatives of LSC, to restrict and to liberate. From 2005 this tension led to some policies that Teresa welcomed and some she considered ill-conceived. Now, in 2009, LSC could shift gears again.

Teresa felt sympathy for LSC. As a manager, Teresa believed in accountability and results. She also knew that creative, aggressive advocacy was not possible if lawyers viewed themselves like government employees. They had to feel both autonomy and that they were

part of an outstanding law firm serving the broader interests of their communities. Consistency and control had to co-exist with risk taking and audacious goals. Compliance with all the rules (of ethics, evidence, and LSC) was necessary; improving clients' lives, however, was the raison d'être.

Acceptance of this duality helped Teresa understand the Polonius-like remarks that her predecessor, Bruce Iwasaki, used to mutter. During the hurried transition period before he left to pursue a career designing costumes for figure skaters (Terry Brooks of the ABA Division of Legal Services, an expert on ice skaters' attire, arranged the job for him), Bruce seemed to belabor the self-evident point that outcome statistics could be used for marketing or used for management, but these functions were quite different. She thought it was obvious that telling the outside world what legal aid did was distinct from revealing to ourselves the impact we had on the outside world. Marketing was used to generate funding and support. The management questions are whether we are trying to do the right things – a priorities issue – and whether we are succeeding – a performance issue.

Bruce suggested that the Legal Services Corporation sometimes conflated (or at least blurred) those goals, and also failed to recognize that the precision and rigor required for these distinct purposes were completely different. But he told Teresa that that was over ten years ago and not to quote him. In any case, Teresa was planning to argue that, based on her experience, there was a third, more important use for outcome measures.

One mission or two

In its thirty-fifth year, LSC was at another critical juncture. Early in 2009 the president of LSC, who had devoted over four decades as a leader in legal services, was appointed to the U.S. Court of Appeals for the Second Circuit. The resulting power vacuum exacerbated the conflict between the two missions of LSC. LSC is a creature of Congress and it has bruises from the numerous life or death struggles on the Hill. With the best survival instincts, one part of LSC's mission must ensure that Congress' directives are followed. The compliance avatar seeks consistency, no surprises and no controversy. The other side of LSC is focused on clients. State planning, client-centered advocacy, and the concern for quality emphasize involvement with other stakeholders to increase the number of low-income people throughout the state who have access to justice. LSC sought to reconcile these two goals in describing its mission: providing high-quality, client-centered civil legal services to the eligible poor in accordance with the mandate of Congress. That statement was surely correct, but there was no question which part of the mission mattered to some elements of Congress, and which was the priority for the field: the falcon cannot hear the falconer.

Teresa reflected on LSC's initiatives concerning outcome measures over the past five years. Some were successful and some were borderline disastrous. The most problematic was when LSC decided in 2005 to impose universal outcome measures that grantees reported to LSC. LSC required programs to count the number of cases that resulted in increased safety and security for the client, and to obtain surveys of client satisfaction. It hoped for, and often received, data that was useful in making the case that federal funding of civil legal services barely dented the burgeoning need. The marketing function was thus usefully fulfilled.

But requiring these statistics also caused the audit side of the LSC brain to swing into action. Anytime you require a grantee to count something, this viewpoint maintains, it has got to be verifiable. Despite initial assurances that this would not be like the CSR fiasco, and would not impose undue burdens on grantees, it caused a big mess.

Dangers of asking for numbers

Asking about clients' safety and security seemed innocuous enough, but things are not so simple. One can win a lawsuit, even obtain money and reform the law, and the client is still poor, unemployed and unable to afford decent housing. On the other hand, one can lose a lawsuit and, because of publicity or pressure, change the system. Aside from that, safety and security were such supple terms that the natural inclination was to interpret them broadly.

And client satisfaction surveys were of very limited use. It was difficult for recipients to get a scientifically random or even a representative sample. And when asked what services they wanted, clients invariably listed their personal problems. This did little to help identify issues the program might be overlooking.

LSC began to see some of these difficulties when it saw how varied the survey forms were from program to program. When it tried to impose some consistency on them however, it took a year to thrash out the wording for the form. Then programs translated these surveys into different languages, but because the translations varied, a comparison among them was infeasible. The whole point of having a single survey instrument was undermined.

What then began to happen was like an urban myth. LSC was perceived as favoring programs on their survey response rate, no matter what the client surveys said, which encouraged programs either to churn out survey results or to serve more educated, English-speaking clients. So, not only was energy spent by programs capturing meaningless data – shades of CSRs – LSC spent time and money verifying the accuracy of dubious numbers. And the effect on the field was to skew priorities and constrict access in unfortunate ways.

LAFLA tried many ways to get feedback on quality. It had been using surveys in various ways for many years. Some substantive units sent questionnaires to all clients, or phoned them, or had simple questionnaires in the waiting room. Teresa was always gratified to read the favorable comments that applicants for service submitted about how they were treated. (She shared the most appreciative notes with her board.) But she was disturbed at how often clients' feedback bore no relationship to what Teresa considered the actual outcome of the case. Indeed, she found that positive client surveys usually reflected how energetic and welcoming the receptionists were, and while that was important, she wondered whether a doctor who had a wonderful bedside manner but did not cure patients should keep her license.

Multiplying outcome codes

LAFLA also used outcome codes for cases – 148 of them – way too many Teresa thought, but she knew how committees worked and what happens when lawyers get in a room and list different results for different types of cases. While consistent capturing of these outcomes was spotty at first, by late 2004 LAFLA was regularly reviewing and sharing the data back to the front line staff. The problem was that as funding diversified, each funder wanted its unique outcomes to be reported. The county funding domestic violence work, for instance, demanded to know the number of cases in which LAFLA appeared in court once, and in two hearings, and in three hearings, and in four or more. Utterly useless information, but you couldn't get paid without it. That made the list even more unwieldy.

LAFLA also ran focus groups of recently-served clients, both extended service and brief service. The clients enjoyed the sessions because they were so seldom asked for their opinion. Staff learned how much their work was appreciated, which itself was a valuable reason to conduct the focus groups. Staff also discovered the paramount value that many clients place on waiting times, being remembered, and courtesy. Clients also stated that the assistance they received in advance of their self-represented pro se appearances helped them be better prepared. On the other hand, the focus groups shed no light on professional skill of the advocates they encountered. Sometimes the focus groups yielded information that could be used in advocacy. For example, a few years ago clients reported that the court clerks gave out erroneous information. As a result, LAFLA was able to correct the errors.

In late-2006 the Republican Latino caucus of Congress objected that LSC's survey procedure discriminated against non-English speakers. This charge was certainly untrue as a general matter, but it was bad publicity. Then the General Accounting Office whipsawed LSC. This time it was not because grantees had mischaracterized or miscounted outcomes, but because GAO concluded that auditing outcome statistics was a waste of resources. LSC changed course.

Stories of enduring results

By 2007 LSC adopted a modified version of what Bruce had proposed at a meeting in Cincinnati in 2004. Then some people thought Bruce had eaten too much goetta. While Bruce opposed LSC requiring any particular outcome measures for all grantees, he argued that if it was going to do so, LSC should urge programs to capture stories and report data on the long term community results of impact advocacy.

Bruce's point was that the focus on quality should relate to and evolve from state planning. He argued that outcomes should be tied to the ABA Standards for Providers of Civil Legal Services to the Poor, which states that "a legal services provider should strive to achieve lasting results responsive to client identified needs and objectives." (Standard 6.5) This includes improving laws and practices that affect clients: "In some cases, this may require litigation that raises either statutory or constitutional questions; administrative representation that seeks change in agency rules, regulations and practices of general application; or legislative representation that seeks statutory change." As the commentary noted, broad challenges to legal problems confronting clients "may be the most cost-efficient way to utilize the limited resources available to meet the legal needs of the poor."

Link to state planning

Eventually, people took the core of Bruce's point and more articulately persuaded LSC that the issue of quality could not be divorced from LSC's long standing effort to build thriving state justice communities. They suggested that otherwise LSC would appear to have simply switched to a new campaign rather than build on one in which many had invested a great deal. (There was enough whining in the field.) If building broad-based coalitions to serve the legal needs of all low income people in a state was a good thing for LSC programs to do, then they should report on how those collaborations resulted in lasting impact. To evaluate client-centered advocacy, it was urged, LSC should receive reports on advocacy that alters for the better the environment and systems – the housing market, the education system, the hospitals – low income people depended on.

In 2007 LSC asked its grantees to describe efforts to achieve lasting results in responding to clients' needs. The reports were intended to show whether a comprehensive, integrated, client-centered effort to improve client lives had succeeded. The reports defined large scale community issues, described strategies and collaborations used to address them, and explained the results and lessons learned. The reports also included how the program used technology and implemented staff development and training.

Emphasizing lasting results meant that the focus was on quality more than numbers, but numbers came into play. Examples were the number of people who not only left welfare but escaped poverty, dollars won for clients in suits against slumlords, number of people who obtained employment because job barriers were removed, and the increase in clinic visits resulting from advocacy to expand health care services.

In part because some in LSC were not sure they wanted to know these numbers, LSC stopped requiring them, and thus stopped auditing them. Narratives of good results began to supplant arbitrary statistics, and LSC found that for marketing purposes – to present to members of Congress – compelling success stories were far more valuable anyway. LSC learned that the more numbers it required of the field, the more problems it bought for itself; the more it scoured the field for terrific stories of legal aid helping people, the better relations it had on the Hill.

Actually, Teresa had no problem with capturing outcome numbers. If anything LAFLA had too many outcome categories. But she insisted that the firm either put the outcome data to use or stop collecting it. She began posting monthly reports and graphs in every office. Staff were able to compare results over time and discuss what influenced the trends.

One result was a surprise to everyone. Whereas the required numbers on clients achieving stability resulted in unreliable statistics and a sticky GAO investigation, measuring lasting results actually won bipartisan political support.

Elected officials of all political persuasions want the government to run efficiently. When LSC-funded programs were able to show that some government agencies were

inefficient, wasting money, and violating the law, and had data to back up their claims, that was welcomed by legislative committees.

In addition, large foundations and regional funders appreciated the shift toward effecting systematic change and began asking for similar data in requests for funding proposals. The upshot was more bipartisan support for Legal Services, greater collaborations between Legal Services programs and local officials, government agencies, community advocates, courts, and other key stakeholders. It also opened avenues for greater funding beyond the Federal government.

Outcome numbers used as weapons

Teresa thought about this when she started outlining what she would write for the 2009 conference. As Bruce told her back in 2005, measuring relevant outcomes helped in (1) evaluating and managing work and (2) publicizing community needs and the value of Legal Aid's services. But there was a third purpose that she felt was even more important.

Capturing the effects of advocacy could be used not only to measure, but to achieve advocacy goals. This was particularly true for impact litigation and policy advocacy. Once LAFLA's lawyers saw the power of aggregating their daily encounters to persuasive effect, the more enthusiastic they became about capturing outcomes.

For instance, LAFLA's Eviction Unit had routinely captured outcomes of not only the cases they litigated, but for those clients who were advised and appeared pro se. In 2003, their first year of counting, the eviction unit won 94% of its cases, obtained over $1.2 million in waived rent and relocation costs, and secured over 20,000 extra days of shelter – almost 55 years – for those moving. Because of Legal Aid's presence in the courthouse, the quality of the pleadings, and the preparation given to clients, LAFLA dramatically increased the success rate of pro se litigants from dead zero to over 30 percent. These were impressive numbers that were perfect for marketing,

But Teresa got the eviction unit not only to collect the data, but use it in a more strategic way. LAFLA began collating the outcome of different judges sitting in the eviction court. That led to graphs showing the judges who were most hostile to unrepresented low-income tenants. The firm presented that data to the presiding judge. Without having to make it public, LAFLA began to see noticeable improvement in the quality and fairness of those judges. Now the family law unit is considering this approach on a broader scale, comparing results in different court houses.

The same idea was applied to the error rate at various welfare offices. With some help from academics, a testing organization, and law students, LAFLA focused on four welfare offices and demonstrated clear differences among them in providing equal access to services. At two offices testers found that applicants were turned away without a chance to submit forms. The power of having this data was that Legal Aid knew more about the welfare department's operations than the welfare department's own management. As a result, the government benefits unit got both a contract to train welfare department staff, and a consent decree against violations based on language and disability. By using data collected from ongoing individual advocacy and analyzing them, the firm was able to target

those performing below par. Thus, the effect was to raise the average. In Lake Wobegon fashion, where all the children are above average, the goal was to continue targeting those at the bottom and continue nibbling up.

The 2009 conference is going to be in the same city the 2004 conference was. Teresa did not know much about Cincinnati, other than Charles Keating, Pete Rose, Marge Schott, Jerry Springer and Larry Flynt. Bruce had told her that one of his very favorite legal aid executive directors was Mary Asbury, but Mary was in Nancy, the capital of the French province of Lorraine, and the sister city of Cincinnati. Mary was on a city-sponsored delegation studying power generation. She had attended the 2004 conference, but deeply regretted that she had to be in France instead of attending the meeting this year.

Teresa will propose that outcomes be used not only to manage and market, but as an instrument of advocacy. An outcome, she would argue, was not necessarily the end. Outcomes could be inputs. Data on what happens to clients could become a sword – like an expert witness or a well-crafted legal theory – to change systems that inherit and oppress the poor.

Bruce G. Iwasaki
May 2004